Dario Gjorgjevski

Curriculum vitæ

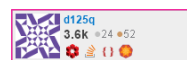**@** dario.gjorgjevski@gmail.com

✉ Alfred-Jung-Straße 14
10369 Berlin, DE

📱 +49 15 25 8 66 68 37

🔑 $ FPR='744A4F0B4F1C9371'
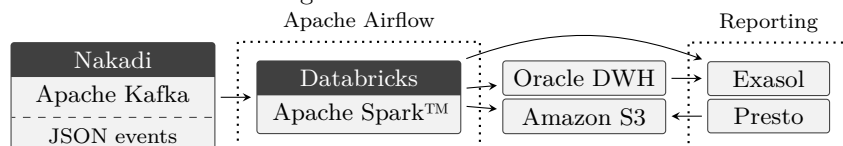$ gpg --recv-keys "$FPR"

⌨ /d125q

1 Experience

Data Engineer — 2019–09/
Zalando SE — Berlin, DE

Part of the team handling Zalando SE's *core data*: articles and sales.

Data Engineer — 2018–07/2019–09
SO1 GmbH — Berlin, DE

- Responsible for a Vertica DWH running on Microsoft Azure:
  - Defined *data vault* and *dimensional* models of client data.
  - Developed ETL processes to ingest data from Blob Storage, Apache Kafka®, and SFTP servers.
  - Wrote and optimized queries to monitor KPIs and to compute features for machine learning.
  - Implemented a UDx to evaluate LightGBM models directly inside the DWH.
- Translated business rules to *minimum cost maximum flow* problems, yielding 15 % greater value than the previous greedy algorithms.
- Developed a microservice for top-k nearest neighbor queries in real time by utilizing *locality-sensitive hashing* with *MinHash* signatures.

Data Scientist — 2017–11/2018–06
Infinite Analytics, Inc. — Skopje, MK

- Scraped clients' websites using Scrapy.
- Developed an Apache Spark™ application to compute and visualize *actionable insights* using over 2 billion facts about 50 million customers.

Research Intern — 2017–07/09
EPFL, LCA2 — Lausanne, CH

- Implemented software agents with `asyncio` and Mininet in T-RECS.
- Modeled smart grid power traces at a timescale of 20 ms using approaches based on wavelets and long-range dependence.
- Increased the resolution of mean-aggregated measurements using deep learning for super-resolution.

Research Intern — 2016–06/09
EPFL, LASEC — Lausanne, CH

Studied and improved upon the complexity of state-of-the-art solving algorithms for the *Learning With Errors* (LWE) problem.

2 Strengths

2.1 Theoretical Knowledge

Algorithms, Data Structures	Probability Theory, Statistics
Databases	Distributed Systems, Big Data Technologies
Linear Algebra	Cryptography

2.2 Hands-On Experience

Python	GNU/Linux
SQL incl. PL/SQL, Apache Spark™	R incl. `Tidyverse` and `data.table`
Redis, Apache Kafka	Apache Airflow
Computer Algebra	LaTeX incl. Ti*k*Z/PGF

- Proficient in both *scientific computing* and *software development*; and capable of writing clean, well-documented, and scalable code.
- Experience with the *agile methodology*, Git for (distributed) version control, and Jenkins for CI/CD.
- Strong academic exposure to programming language theory along with various Lisps, Haskell, and Standard ML.

3 Honors

- Graduated *summa cum laude* with a perfect GPA from the Ss. Cyril and Methodius University.

- Best student paper at the *14th International Conference on Informatics and Information Technologies* held on Apr. 7–9, 2017 in Mavrovo, Macedonia.

- Scholarships to attend the 2016 and 2017 editions of the *Summer School on Real-World Crypto and Privacy* held in Šibenik, Croatia.

- Dean's list at the Ss. Cyril and Methodius University and merit-based scholarships — awarded to the top 2.5 % students on a national level — throughout the entire duration of study.

4 Education

Computer Science & Engineering (B.Sc.)   2013-09-15/2018-01-30
Ss. Cyril and Methodius University   Skopje, MK
GPA of 10.00; scale from 5 (E/F) to 10 (A).
Thesis: "Error-Correcting Codes in the Rank Metric" [2].
Publications: [1, 3].
As a senior, conducted computational exercises and examinations as
well as homework assignments in:

Linear Algebra   Least squares, linear codes, and low-rank approxima-
tions in SageMath and Mathematica®.

Statistics   Data visualization, Monte Carlo methods, inference,
hypothesis testing, and linear regression in R.

Databases   ER models, relational algebra, and ANSI SQL.

Presentations available at /d125q/Presentations.

Over 40 *Massive Open Online Courses* on topics related to game the-
ory, probabilistic graphical models, Bayesian statistics, combinatorics,
automata and formal languages, mathematical optimization, etc.
- -
Certifications available at /d125q/Personal/tree/master/Certifications.

---

5 Theses

[2] **Dario Gjorgjevski**. "Error-Correcting Codes in the Rank Metric." With
Applications to Cryptography. Bachelor's Thesis. Under sup. of Simona
Samardjiska. Ss. Cyril and Methodius University, Jan. 24, 2018. eprint:
http://diplomski.finki.ukim.mk/Upload/PublicFile/1814.

6 Publications

[1] **Dario Gjorgjevski**. "Combining LWE-Solving Algorithms." In: *Proceed-
ings of the 14th International Conference on Informatics and Information
Technologies* (Hotel Bistra, Mavrovo, Macedonia, Apr. 7–9, 2017). Ed. by
Aleksandra Popovska–Mitrovikj, Biljana Tojtovska, and Kire Trivodaliev.
2017, pp. 165–170. ISBN: 978-608-4699-07-1. eprint: http://ciit.finki.
ukim.mk/data/papers/CIIT2017.pdf.

[3] **Dario Gjorgjevski** and Dejan Gjorgjevikj. "Using Distributed Representa-
tions to Identify Genders and Age Groups of Twitter Users." In: *Proceedings
of the 15th International Conference on Informatics and Information Tech-
nologies* (Hotel Bistra, Mavrovo, Macedonia, Apr. 20–22, 2018). Ed. by
Nataša Ilievska and Georgina Mirčeva. 2018, pp. 2–7. ISBN: 978-608-4699-
08-8. eprint: http://ciit.finki.ukim.mk/data/papers/CIIT2018.pdf.

7 Projects

C-like language ⇒ PostScript transpiler
Transpiler implemented in Flex and GNU Bison to
translate a C-like language for *turtle graphics* to
PostScript.

Trusted timestamping
Flask application for a simple file-sharing service
which also provides *trusted timestamps* as specified
in RFC3161 and implemented in OpenSSL.

AS-level robustness of the Internet over time
Simulation of random and targeted attacks against
the Internet topology. Notebook and source code
available at /d125q/Internet_Robustness.

Predicting readmission of diabetic patients
- Learning from imbalanced data using `mlr`.
- *Fully reproducible reporting* using LATEX, `mlr`,
  and `knitr`.

Available at /d125q/Diabetic_Patients.

Survey of the MinRank problem
SageMath implementations of:
- Algorithms for solving MinRank; and a
- Zero-knowledge authentication protocol based
  on MinRank.

Available at /d125q/MinRank.

Checksum verification on LPC1769
C program to verify MD5 checksums of payloads
stored on an LPC1769 microcontroller. Won prize
for highest throughput. Source code available at
 /d125q/DMA_Workshop.

Substitution ciphers for Macedonian text
- Create substitution ciphers; and
- Break substitution ciphers using Markov chain
  Monte Carlo (MCMC) methods based on uni-
  gram and bigram frequencies.

Mathematica® notebook and corpus with Macedo-
nian text suitable for frequency analysis available at
 /d125q/Macedonian_Substitution_Ciphers.